

Dogs' Gaze Following Is Tuned to Human Communicative Signals

Ernő Téglás,^{1,2} Anna Gergely,³ Krisztina Kupán,² Ádám Miklósi,³ and József Topál^{2,*}

¹Cognitive Development Center, Central European University, Hattyú 14, 1051 Budapest, Hungary

²Comparative Behavioural Research Group, Research Institute for Psychology, Hungarian Academy of Sciences, Victor Hugo Street, 18-22, 1132 Budapest, Hungary

³Department of Ethology, Loránd Eötvös University, Pázmány P. 1c, 1117 Budapest, Hungary

Summary

Recent evidence suggests that preverbal infants' gaze following can be triggered only if an actor's head turn is preceded by the expression of communicative intent [1]. Such connectedness between ostensive and referential signals may be uniquely human, enabling infants to effectively respond to referential communication directed to them. In the light of increasing evidence of dogs' social communicative skills [2], an intriguing question is whether dogs' responsiveness to human directional gestures [3] is associated with the situational context in an infant-like manner. Borrowing a method used in infant studies [1], dogs watched video presentations of a human actor turning toward one of two objects, and their eye-gaze patterns were recorded with an eye tracker. Results show a higher tendency of gaze following in dogs when the human's head turning was preceded by the expression of communicative intent (direct gaze, addressing). This is the first evidence to show that (1) eye-tracking techniques can be used for studying dogs' social skills and (2) the exploitation of human gaze cues depends on the communicatively relevant pattern of ostensive and referential signals in dogs. Our findings give further support to the existence of a functionally infant-analog social competence in this species.

Results

Although recent research has provided important evidence about dogs' social communicative skills [2], it is still unclear whether dogs' gaze following is tuned to cues that signal the human's communicative intent (e.g., eye contact, verbal addressing). In order to investigate this first, we have collected eye-gaze data from 16 adult, task-naïve pet dogs. Subjects were presented with a series of movies in which a human female turned her attention toward one of two identical containers either in an ostensive-communicative (O) or in a nonostensive (NO) manner. In the ostensive condition, the human actor overtly expressed her communicative intent, whereas in the nonostensive condition, we removed the ostensive signal from the stimulus. Each trial consisted of introductory, addressing, and cueing phases (Figures 1A–1C). All dogs participated in both O and NO trials that were different only in the addressing phase.

We obtained valid data for analysis from 13 dogs in the ostensive and 14 dogs in the nonostensive condition; however, only 11 of them provided valid data in each of these conditions (for the validity criteria see [Data Analysis](#)). In the addressing phase, dogs spent similar amounts of time gazing toward the human actor in the two conditions (mean $\pm$ SEM: 1,088.8 $\pm$ 181.1 ms in O and 980.9 $\pm$ 267.8 ms in NO conditions, ns) and invested a comparable amount of time scanning the region containing the actor's face relative to the whole body: mean $\pm$ SEM = 0.46 $\pm$ 0.09 in the O condition and 0.55 $\pm$ 0.10 in the NO condition [paired *t* test *t*(10) = −0.88; *p* = 0.39] showing that in the addressing phase, the human actor evoked the same level of visual attention in both conditions.

Next we analyzed whether dogs looked longer at the gaze-congruent area ([Figure 1D](#)) as compared to the gaze-incongruent area (cumulative accuracy). In accordance with infant eye-tracking studies (e.g., [1]), difference scores were calculated for this variable. We found that subjects looked longer to the gaze-congruent area than to the gaze-incongruent area after having seen ostensive addressing (one sample *t* test, *t*(12) = 2.382; *p* = 0.034). However, this was not the case for the nonostensive condition in which the difference score did not differ from zero [one-sample *t* test, *t*(13) = −0.756; *p* = 0.46], indicating no tendency to follow the human's gaze in the absence of communicative addressing ([Figure 2](#)). This differential sensitivity to human referential gestures is strikingly similar to that found in a study of 6.5-month-old human infants [1]. A similar analysis on dogs' first look did not reveal any significant bias toward the gaze-congruent area in O or in the NO conditions [one-sample *t* test, *t*(12) = 1.167; *p* = 0.266; *t*(13) = −0.105; *p* = 0.91, [Figure 2](#)].

A within subject analysis of the difference scores for cumulative looking time in the two experimental conditions was run on the 11 subjects that gave valid data in both conditions. This analysis shows that dogs were more likely to follow the model's gaze in a gaze-congruent manner in the O than in the NO condition [*t*(10) = 2.49; *p* = 0.03; the effect was independent of presentation order; see [Supplemental Results](#) available online]. However, no difference was found between conditions for the first look [*t*(10) = −1.21; *p* = 0.25].

We also explored the spatiotemporal pattern of eye movements during cueing phase, investigating how the gaze points move away from the midline of the display and approach the target objects. Gaze points were averaged into 1 s bins and were projected to the *x* axis of the display ([Figure 3](#)). The averaged eye movements showed a greater proximity to the target object only in the ostensive condition.

Thus, we may conclude that dogs' context-dependent responsiveness to human head turning mirrors the specific effect of human ostensive communication on dogs' cognitive processing. Although the stimuli in the two conditions were equally successful in orienting dogs' attention toward the actor's head in the addressing phase, only the ostensive cues led to gaze following. However, in order to elicit a comparable saliency of the addressing phases in the two conditions in the NO condition, we displayed a moving attention-getter on the model's forehead ([Figure 1C](#)) similar to the infant

*Correspondence: topaljozsef@gmail.com

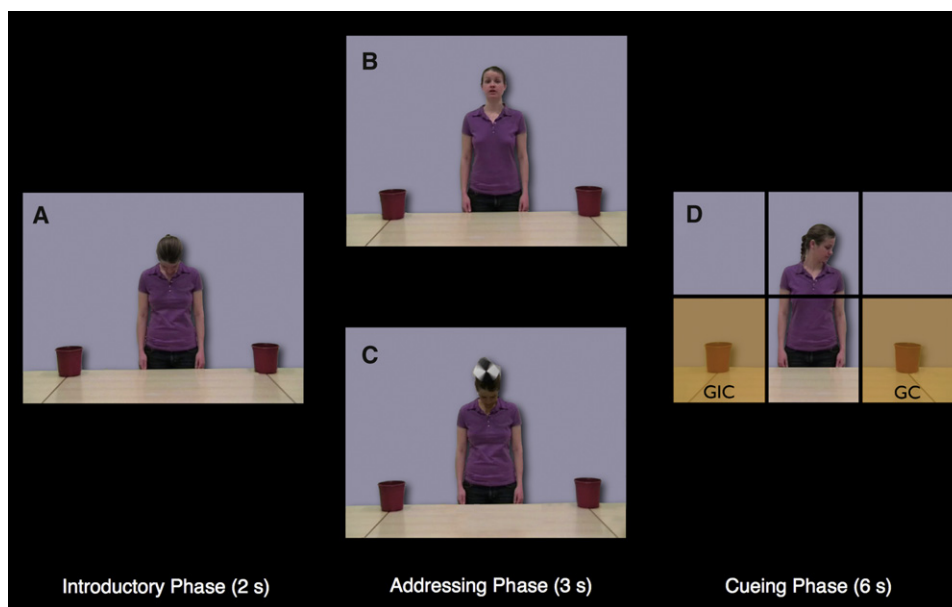


Figure 1. Selected Frames from the Stimuli in the Ostensive and Nonostensive Conditions

Each trial presented video recordings that started with an introductory phase during which the model that had two pots on each side was facing down in a still position for 2 s (A). The second phase was an addressing phase that lasted for 3 s and differed according to the experimental conditions. In the O (ostensive; A, B, D) condition, the model raised her head, looked straight at the dog, and addressed the subject ("Hi dog!") in a high-pitched voice (B). In the NO (nonostensive; A, C, D) condition, with her head facing down, the model addressed the dog using low-pitched voice ("Hi dog!") while a salient moving image was overlaid on the head (C). This attention-getter was present for 2 s and served to create attentional demand similar to that in the ostensive condition. The verbal signal in the two conditions had similar duration and intensity but differed in pitch. In the cueing phase (6 s), the model turned her head toward one of the two containers (1 s) and remained motionless (5 s) while showing neutral facial expression (D). GC indicates the gaze congruent and GIC the gaze incongruent regions of interest (ROI).

study [1]. Importantly, this raises the possibility that not the absence of ostensive cues but the artificial nature of this stimulus has contributed to the reduced gaze following in the NO condition.

To exclude this, we measured the gaze-following behavior of 13 additional experimentally naive dogs in a baseline control condition (BC) in which (1) the human actor turned her head without providing any ostensive cues (eye contact and addressing), (2) the salient attention-getter was removed from the addressing phase, and (3) the verbal addressing was replaced with neutral beep sound of similar duration and intensity in order to keep the auditory marking of this

sequence comparable, while attracting the dogs' attention to the screen.

The data showed that gazing toward the region containing the actor's face relative to the whole body in the addressing phase was comparable to that found in O and NO conditions (mean $\pm$ SEM, 0.60 ± 0.07). However, dogs looked longer toward the body of the protagonist in both O and NO conditions than in BC [400.2 ± 106.9 ms; two-sample *t* tests: O versus BC $t(21) = -3.44$, $p = 0.002$; NO versus BC $t(21) = -2.187$, $p = 0.04$]. Thus, the combination of visual and audio components of the stimuli available in the addressing phase of O and NO conditions (direct gaze and infant-directed

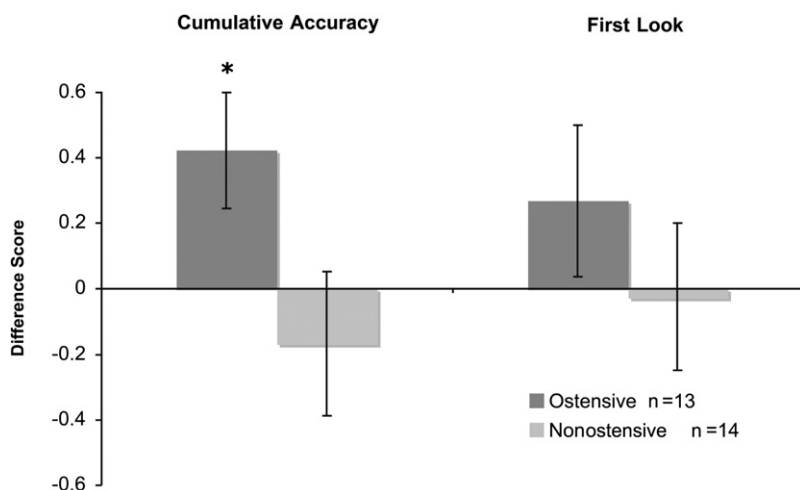


Figure 2. Difference Scores Calculated in O and NO Conditions for Different Measures of Gaze Following

* $p < 0.05$; error bars represent SEM.

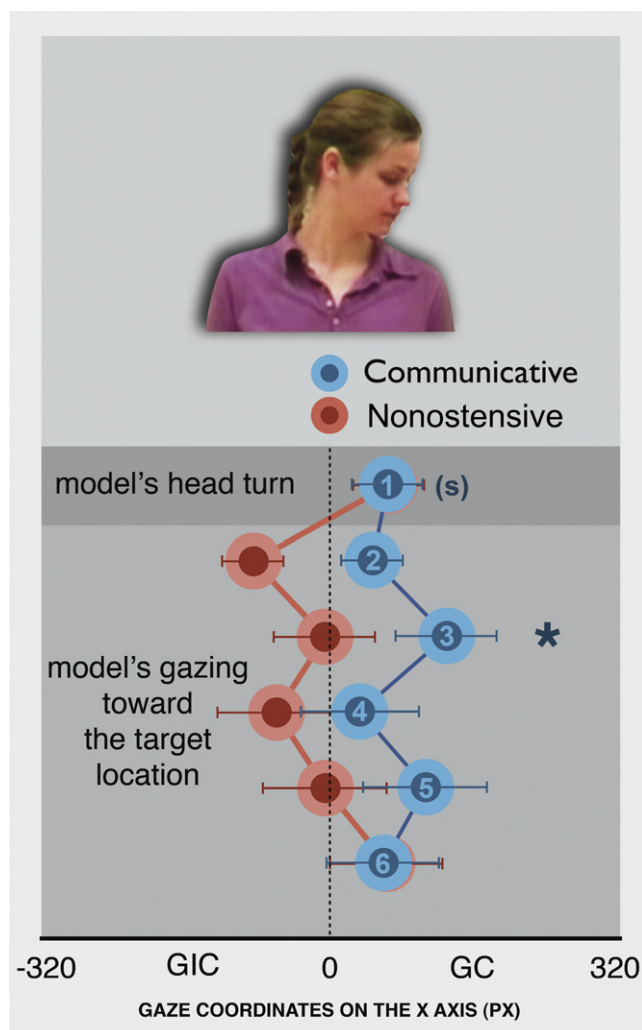


Figure 3. The Temporal Dynamic of the Eye Movements during the Cueing Phase in O and NO Conditions

The gaze data recorded from the lower half of the screen capture the main trends of eye movements (with gaze coordinates projected to x axis; resolution $X = 1,280$ pixels) as the mean gaze points move away from the midline of the display toward the GC or GIC region. After the actor's head movement, there is a peak that differs significantly from the central axis of the display only during the ostensive condition (* $p < 0.05$; error bars represent SEM, 1° visual angle is approximately 40 pixels; y axis represents time).

speech or visual grabber and adult-directed speech) attracted more attention toward the human actor. Importantly, however, the accuracy indexes calculated for the cueing phase did not capture gaze following in BC [mean $\pm$ SEM of cumulative accuracy: -0.06 ± 0.14 ; one-sample t test, $t(12) = -0.433$; $p = 0.67$; first look: 0.11 ± 0.16 ; one-sample t test, $t(12) = 0.695$; $p = 0.5$]. Furthermore, dogs followed the actor's gaze significantly less in the BC compared to the O condition [cumulative accuracy, two-sample t test, $t(24) = -2.107$, $p = 0.045$]. These data suggest that the lack of gaze following, when there are no ostensive signals, cannot be accounted for by the confounding effect of "artificial" salient attention getter used in the NO condition, also providing further support for the significant role of ostensive signals in dogs' gaze response.

Discussion

Our results indicate striking similarities between adult pet dogs and preverbal infants [1] regarding their context-specific responsiveness to human referential signals. This is supported by the dogs' bias to look longer toward the gaze-congruent area in the cueing phase of the ostensive conditions, but not in the non-ostensive conditions. However, first-look measures did not show significant context-specific differences despite the fact that this test variable is usually reported to be one of the strongest indexes to capture human gaze following in presence of ostensive-communicative cues. The discrepancy between the two measures seems to suggest that although being overall sensitive to the ostensive signals, dogs, in contrast to 6.5-month-old infants [1], might be less responsive to the actual onset of these cues. At an earlier age, even human infants show similar patterns: although they generally prefer to fixate on gaze-congruent objects, this bias is not reflected in their first gaze shifts [4]. In line with data from a recent study [5], one may also argue that dogs are generally less accurate in their first fixations. Finally, it is also possible that our subjects, by being allowed to move freely, might have produced discontinuous eye-gaze recording that introduced more noise in the first-look analysis.

Gaze-following behavior among humans is an early emerging pervasive response [6] and is frequently considered as a window into social cognition of different nonhuman species [7]. For instance, dogs have a robust ability to share attention with humans, and they are very skillful in using human gaze in object-choice situations [8]. Dogs are sensitive to the direction of human visual attention [9] and are skillful users of human directional signals that have potential referential significance [10]. Moreover, increasing evidence suggests that dogs show early [11] and infant-like sensitivity [12] to cues that signal the human's communicative intent. These findings raise the possibility that, independently from the actual underlying mental mechanisms, dogs display analog functioning in terms of performance to preverbal infants in ostensive-communicative situations. Besides the susceptibility to human ostensive cues, dogs' gaze-following behavior may also be considered as a socially facilitated orientation response with aspects linked to associative understanding of the net utility of the co-orientation with others.

It is also important to note that eye contact was used in combination with ostensive addressing in our study; thus, the question about the contribution of individual ostensive cues to dogs' tendency to follow human gaze is still unanswered. Whether these cues act independently or in combination and whether they can be ranked according to their efficiency in eliciting the communicative understanding of certain social interactions should be targets for further investigations.

These results provide further support for the notion that dogs might have evolved a special, functionally infant-analog "cognitive mindset," which facilitates the emergence of communicative interaction with people as a result of proper socialization to human environments. Such socially motivated "cue-driven" gaze following displayed by preverbal infants and dogs is a necessary but probably not a sufficient precondition for a deeper understanding of the intentional-communicative nature of referential signals.

Experimental Procedures

Subjects

Sixty-one adult pet dogs were tested at the Department of Ethology, Eötvös University. Eye-gaze data of 29 dogs (16 participated in O and NO conditions and 13 in BC condition) were considered for analysis (see

Supplemental Experimental Procedures). Thirty-two dogs were excluded because of inattentiveness (five dogs in O and NO and seven dogs in BC conditions) or unsuccessful calibration (20 dogs).

Apparatus

The gaze data was collected at 50 Hz by a Tobii X50 Eye Tracker (Stockholm, Sweden). The eye tracker had 0.5–0.7 degree accuracy, 30 × 16 × 20 cm freedom of head movement. The stimuli were presented on a 17-inch LCD monitor positioned behind the eye tracker. The owner made the dog stand, sit, or lie down in order to get optimal eye-gaze data (at a distance of approximately 60 cm). The owner sat behind the dog and nodded his/her head while looking down and avoiding verbal interactions.

Procedure

Warm-Up Trials

Before the experiment, dogs participated in ten trial warm-up sessions that served to elicit the dog's interest in the plastic pots shown in the experimental trials (see Supplemental Experimental Procedures).

Data Acquisition

The eye-gaze recording was preceded by a five-point calibration phase following the infant calibration protocol of Clearview 2.5.1. Same software presented the video clips in the test trials. Dogs participated in the test trials only if contributed to minimum four calibration points.

Test Trials

Dogs that were exposed to both O and NO conditions received two blocks of six trials, where the blocks were composed of either video materials containing ostensive cues (Figures 1A, 1B, and 1D) or nonostensive cues (Figures 1A, 1C, and 1D). The blocks were presented in counterbalanced order, and the direction of the model's gaze was also counterbalanced. Subjects assigned to baseline control group received only one block of six trials (Figures 1A and 1D, but without attention-getter in phase C). The study was approved by the Animal Care and Experimentation Committee, Eötvös University, and the research was performed in accordance with the national regulations on animal experimentation.

Data Analysis

The screen was divided on the horizontal axis in three equal areas: left, middle, and right (see Figure 1D). Our statistical analysis was based on the eye gaze collected from the lower half of the lateral (left and right) areas (10.5° × 12° visual angle) that were defined as regions of interest (ROI) during the cueing phase. Trials were accepted as valid only if they provided more than 200 ms eye-gaze data from the ROI. These criteria were implemented in order to exclude short transitions of the gaze that just happened to pass the ROIs. Although during the cueing phase eye gaze was recorded from the regions of interest in 69% of the trials, taking these criteria into consideration, dogs provided 27% valid trials in the O and 34% in the NO conditions. For dogs that participated in BC condition, eye gaze was recorded from the target regions of interest in 60% of the trials, out of which 35% were valid trials.

Scoring

The gaze following was tested along two measures: cumulative accuracy and first look toward the gaze-congruent object. These indexes correspond to those presented by Senju and Csibra [1]. For each of these measures, difference scores were calculated. For instance, trials in which dogs looked only to the side congruent with the model's gaze were coded as correct (c), whereas if the dog didn't look at the correct side, the trial was coded as incorrect (i). When dogs looked at both sides, the trial was classified according to the longer look. Thus, the difference score (d) for the cumulative accuracy was calculated by subtracting the incorrect from the correct trials and dividing the result by the total number of trials where the participant provided valid ROI data [$d = (c-i)/(c+i)$]. The first look was analyzed in a similar fashion, but instead of the time spent in one or the other ROI, only the first gaze record toward these ROIs was considered.

Supplemental Information

Supplemental Information includes Supplemental Results and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.cub.2011.12.018.

Acknowledgments

This work was funded by the Hungarian Science Foundation (OTKA K76043 & NK83997), the European Union (EU FP7 ICT-215554 LIREC), and the Marie Curie Research Training Network.

Received: February 15, 2011

Revised: November 14, 2011

Accepted: December 1, 2011

Published online: January 5, 2012

References

1. Senju, A., and Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Curr. Biol.* 18, 668–671.
2. Topál, J., Miklósi, Á., Gácsi, M., Dóka, A., Pongrácz, P., Kubinyi, E., Virányi, Z., and Csányi, V. (2009). The dog as a model for understanding human social behavior. *Adv. S. Anim. Behav.* 39, 71–116.
3. Lakatos, G., Soproni, K., Dóka, A., and Miklósi, Á. (2009). A comparative approach to dogs' (*Canis familiaris*) and human infants' comprehension of various forms of pointing gestures. *Anim. Cogn.* 12, 621–631.
4. Gredeback, G., Theuring, C., Hauf, P., and Kenward, B. (2008). The microstructure of infants' gaze as they view adult shifts in overt attention. *Infancy* 13, 533–543.
5. Williams, F.J., Mills, D.S., and Guo, K. (2011). Development of a head-mounted, eye-tracking system for dogs. *J. Neurosci. Methods* 194, 259–265.
6. Butterworth, G. (1991). The ontogeny and phylogeny of joint visual attention. In *Natural Theories of Mind*, A. Whiten, ed. (Oxford: Basil Blackwell), pp. 223–232.
7. Emery, N.J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neurosci. Biobehav. Rev.* 24, 581–604.
8. Kaminski, J. (2009). Dogs (*Canis familiaris*) are adapted to receive human communication. In *Neurobiology of "Umwelt": How Living Beings Perceive the World*, A. Berthoz and Y. Christen, eds. (Springer Verlag Berlin), pp. 103–107.
9. Kaminski, J., Bräuer, J., Call, J., and Tomasello, M. (2009). Domestic dogs are sensitive to a human's perspective. *Behaviour* 146, 979–998.
10. Soproni, K., Miklósi, Á., Topál, J., and Csányi, V. (2001). Comprehension of human communicative signs in pet dogs (*Canis familiaris*). *J. Comp. Psychol.* 115, 122–126.
11. Riedel, J., Schumann, K., Kaminski, J., Call, J., and Tomasello, M. (2008). The early ontogeny of human-dog communication. *Anim. Behav.* 73, 1003–1014.
12. Topál, J., Gergely, G., Erdőhegyi, Á., Csibra, G., and Miklósi, Á. (2009b). Differential sensitivity to human communication in dogs, wolves, and human infants. *Science* 325, 1269–1272.